ISYE 6416 Homework 5

Alex Bukharin

April 2020

1 Nonlinear Regression with Spline

1.1 a)

Using Linear Regression to fit the thermal expansion data, we get a slope of 0.021 and a intercept of 7.38. The fitting error is 10.52.

1.2 b)

We find the optimal lambda to be 0.0 by cross validating on lambda values from 0.0 to 1.0 with 0.01 as the step size. I used leave one out cross validation. The fitting error with a lambda of 0.0 is 0 because all points are fitted exactly. The cross validation loss was approximately 0.75 with lambda equal to 0.

1.3 c)

The predicted coefficient and 400 degrees Kelvin is 17.438.

We have attached the code for pca to our submission.

2 PCA for face recognition

2.1 a)

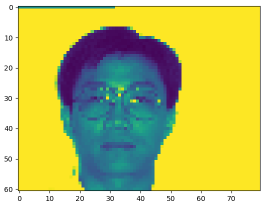

Figure 1: Mean face for subject 14.

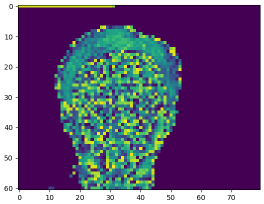

Figure 2: Eigenface 1.

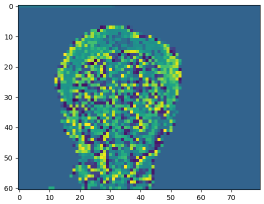

Figure 3: Eigenface 2.

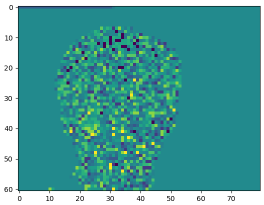

Figure 4: Eigenface 3.

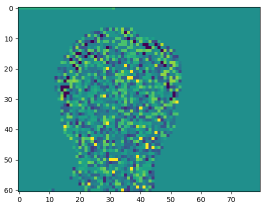

Figure 5: Eigenface 4.

2.2 b)

Using the prediction procedure, we can classify that the test image is subject 14. When we project the test image on the first principal component of subject

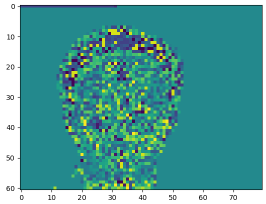

Figure 6: Eigenface 5.

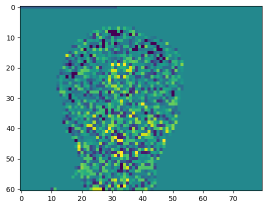

Figure 7: Eigenface 6.

14, we get a projection value of 11,000. When we project the test image on the first principal component of subject 01, we get a value of 7,500.Therefore we can classify the test image as belonging to subject 14.

We have attached the code for pca to our submission.

3 Recommender System

3.1 a)

I followed the suggestions of the instructors and filled in the missing values with the mean. Then I used one-hot encoding to change the categorical data into numerical data. Then I used collaborative filtering to predict the top 5 movies for each user, regardless of whether the movie had been seen or not. The results are included in the attached csv file. I found that the l2 norm generally predicted higher movie rtings than the l1 or l0 norms.

3.2 b)

I followed the same procedure as in section a except with item based recommendation. The results are attached.

3.3 c)

I used the R softImpute package to predict the zeros in the given user matrix. The results are attached as "movie_r.csv".

All of the code for the recommendation system is attached.